```
(5')    1    SDLEQERRAKEKLQEQQ
       18    SDLEQDRLAKEKLQEQQ
       35    SDLEQERLAKEKLQEQQ
       52    SDLEQERRAKEKLQEQQ
       69    SDLEQERRAKEKLQEQQ
       86    SDLEQDRLAKEKLQEQQ          "R"
      103    SDLEQERRAKEKLQEQQ
      120    SDLEQERRAKEKLQEQQ
      137    SDLEQERLAKEKLQEQQ
      154    SDLEQERRAKEKLQEQQ
      171    SDLEQERRAKEKLQEQQ
      188    SDLEQERRAKEKLQEQQ
      205    RDLEQ

      210    RKADTKKNLERKKEHGDILAEDLYGRLEIP
      240    AIELPSENERGYYIPHQSSLPQDNRGNSRD
      270    SKEISIIEKTNRESITTNVEGRRDIHKGHL
      300    EEKKDGSIKPEQKEDKS    316 (3')   (SEQ ID NO:31)
```

FIGURE 1

(5') 1   AAAGCGGATCTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
      52 AAAGCGGATTTAGAACAAGAGATAGACTTGCTAAAGAAAAGTTACAAGAGCAGC
      103 AAAGCGGATTTAGAACAAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAC
      154 AAAGCGGATCTAGAACAAGAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
      205 AAAGCGGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
      256 AAAGCGGATTTAGAACAAGAGATAGACTTGCTAAAGAAAAGTTACAAGAGCAGC
      307 AAAGCGGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
      358 AAAGCGGATTTAGAACAAGAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
      409 AAAGCGGATTTAGAACAAGAGAGACTTGCTAAAGAAAAGTTGCAAGAACAAC
      460 AAAGCGGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
      511 AAAGCGGATTTAGAACAAGAGAGAGACGTGCTAAAGAAAAGTTGCAAGAACAAC
      562 AAAGCGGATTTAGAACAAGAGAGACGTGCTAAAGAAAAGTTGCAAGAGCAGC
      613 AAAGAAGATTTAGAACAA
      630 AGGAAGGCTGATACGAAAAAAATTTAGAAAGAAAAAGGAACATGGAGAT
      681 ATATTAGCAGAGGATTTATATGGTCGTTTAGAAATACCAGCTATAGAACTT
      732 CCATCAGAAAATGAACGTGGATATTATGACAGGAAATATCTCTTTTACCT
      783 CAGGACAACAGAGAGGAATAGTAGAGATTCCAAGGAAATATCTATAATAGAA
      834 AAAACAAATAGAGAATCTATTACAACAAATGTTGAAGGACGAAGGGATATA
      885 CATAAAGGACATCTTGAAGAAAAGAAAGATGGTTCAATAAAAACCAGAACAA
      936 AAAGAAGATAAATCT 950 (3') (SEQ ID NO: 32)

FIGURE 2

RDELFNELLNSVDVNGEVKENILEESQVN&DIFNSLVKSVQQEQQ

HNVEEKVEESVEENDEESVEENVEENVEENDDGSVASSVEESI

ASSVDESIDSSIEENVAPTVEEIVAPTVEEIVAPSVVEKCAPSVE

ESVAPSVEESVAEMLKER   (SEQ ID NO: 24)

FIGURE 3

5' GAA TTC CGT GAT GAA CTT TTT AAT GAA TTA TTA AAT AGT GTA GAT
GTT AAT GGA GAA GTA AAA GAA AAT ATT TTG GAG GAA AGT CAA GTT AAT
GAG GAT ATT TTT AAT AGT TTA GTA AAA AGT GTT CAA CAA GAA CAA CAA
CAC AAT GTT GAA GAA AA AGT TGA AGA AAG TGT AGA AGA AA ATG ACG
AAG AAA GTG TAG AAG AAA ATG TAG AAG AAA ATG TAG AAG AAA ATG
ACG ACG GAA GTG TAG CCT CAA GTG TTG AAG GTA TAG CTT CAA GTG
TTG ATG AAA GTA TAG ATT CAA GTA TTG AAG AAA GTA TAG CTC CAA CTG
TTG AAG AAA TCG TAG CTC CAA CTG TTG AAG AAA TTG TAG CTC CAA GTG
TTG TAG AAA AGT GTG CTC CAA GTG TTG AAG AAA GTG TAG CTC CAA GTG
TTG AAG AAA GTG TAG CTG AAA TGT TGA AGG AAA GGA ATT C 3' (SEQ ID NO: 33)

FIGURE 4

LSA-TER (SEQ ID NO: 23)

NSRDSKEISIIEKTNRESITTNVEGRRDIHK

729S-NRI (SEQ ID NO: 26)
729S-NRII (SEQ ID NO: 27)
729S-Rep (SEQ ID NO: 28)

DELFNELLNSVDVNGEVKENILEESQ
LEESQVNDDIFSNSLVKSVQQEQQHNV
VEKCAPSVEESVAPSVEESVAEMLKER

FIGURE 5

NUCLEOTIDE SEQUENCE OF THE LSA GENE
5' END

(NON-CODING 5' END)

1 AAAGTATACATCTTCCTTCTTTACTTCTTAAA

(CODING 5' END, UNIQUE)

```
33   ATGAAACATATTTTGTACATATCATTTTACTTTATCCTTGTTAATTTATTG
84   ATATTTCATATAAATGGAAAGATAATAAAGAATTCTGAAAAAGATGAAATCA
135  TAAAATCTAACTTGAGAAGTGGTTCTTCAAATTCTAGGAATCGAATAAATGA
186  GGAAAATCACGAGAAGAAACACGTTTTATCTCATAATTCATATGAGAAAACT
237  AAAAATAATGAAAATAATAAATTTTTCGATAAGGATAAAGAGTTAACGATGT
288  CTAATGTAAAAAATGTGTCACAAACAAATTTCAAAAGTCTTTTAAGAAATCT
339  TGGTGTTTCAGAGAATATATTCCTTAAAGAAAATAAATTAAATAAGGAAGGG
390  AAATTAATTGAACACATAATAAATGATGATGACGATAAAAAAAAATATATTA
441  AAGGGCAAGACGAAAACAGACAAGAAGATCTTGAAGAAAAAGCA
```

(CODING 5' END, repetitive)

```
492  GCTAAAGAAAAGTTACAGGGGCAACAAAGCGATTCAGAACAAGAGAGACGT
543  GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACTT
594  GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
645  GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACTT
696  GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
747  GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
798  GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACTT
849  GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTTAGAACAAGATAGACTT
900  GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
951  GCTAAAGAAAGGTTGCAAGAACAACAAAGCGATTTAGA 988   (SEQ ID NO: 34)
```

FIGURE 6

LSA.5'/ATG - -> 1-phase Translation

(SEQ ID NO:36)

DNA sequence 956 b.p. ATGAAACATATT ... AAGGGATTTAGA   linear

(SEQ ID NO:37)
(SEQ ID NO:38)

```
           1 / 1                          31 / 11
ATG AAA CAT ATT TTG TAC ATA TCA TTT TAC   TTT ATC CTT GTT AAT TTA TTG ATA TTT CAT
met lys his ile leu tyr ile ser phe tyr   phe ile leu val asn leu leu ile phe his

          61 / 21                          91 / 31
ATA AAT GGA AAG ATA ATA AAG AAT TCT GAA   AAA GAT GAA ATC ATA AAA TCT AAC TTG AGA
ile asn gly lys ile ile lys asn ser glu   lys asp glu ile ile lys ser asn leu arg

         121 / 41                         151 / 51
AGT GGT TCT TCA AAT TCT AGG AAT TCA ATA   CGA AAT CAC GAG AAG AAA CAC GTT
ser gly ser ser asn ser arg asn ser arg   ile asn his glu lys lys his val

         181 / 61                         211 / 71
TTA TCT CAT AAT TCA TAT GAG AAA ACT       AAA AAT GAA AAT AAT TTT TTC GAT AAG
leu ser his asn ser tyr glu lys thr       lys asn glu asn asn phe phe asp lys

         241 / 81                         271 / 91
GAT AAA GAG TTA ACG ATG TCT AAT GTA AAA   AAT GTG TCA CAA ACA AAT TTC AAA AGT CTT
asp lys glu leu thr met ser asn val lys   asn val ser gln thr asn phe lys ser leu

         301 / 101                        331 / 111
TTA AGA AAT CTT GGT GTT TCA GAG AAT ATA   TTC CTT AAA GAA AAT TTA AAT AAG GAA
leu arg asn leu gly val ser glu asn ile   phe leu lys glu asn leu asn lys glu
```

FIGURE 7A

NUCLEOTIDE SEQUENCE OF THE LSA GENE
3' END

(CODING 3' END, REPETITIVE)

```
1   CAAGAACAACAAAGCGATCTAGAACAAGAGAGACGT
37  GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGATAGACTT
88  GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTTAGAACAAGAGAGACTT
139 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATCTAGAACAAGAGAGACGT
190 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
241 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGATAGACTT
292 GCTAAAGAAAAGTTACAAGAGCAGCAAAGCGATTTAGAACAAGAGAGACGT
343 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
394 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACTT
445 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
496 GCTAAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
547 GCTAAGAAAAGTTGCAAGAACAACAAAGCGATTTAGAACAAGAGAGACGT
598 GCTAAAGAAAAGTTGCAAGAGCAGCAAAGAGATTTAGAACAA
```

(CODING 3' END, UNIQUE)

```
640  AGGAAGGCTGATACGAAAAAAAATTTAGAAAGAAAAAGGAACATGGAGAT
691  ATATTAGCAGAGGATTTATATGGTCGTTTAGAAATACCAGCTATAGAACTT
742  CCATCAGAAAATGAACGTGGATATTATATACCACATCAATCTTCTTTACCT
793  CAGGACAACAGAGGGAATAGTAGAGATTCCAAGGAAATATCTATAATAGAA
844  AAAACAAATAGAGAATCTATTACAACAAATGTTGAAGGACGAAGGGATATA
895  CATAAAGGACATCTTGAAGAAAGAAAGATGGTTCAATAAAACCAGAACAA
946  AAAGAAGATAAATCTGCTGACATACAAAATCATACATTAGAGACAGTAAAT
997  ATTTCTGATGTTAATGATTTTCAAATAAGTAAGTATGAGGATGAAATAAGT
1048 GCTGAATATGACGATTCATTAATAGATGAAGAAGAAGATGATGAAGACT
1099 TAGACGAATTTAAGCCTATTGTGCAATATGACAATTTCCAAGATGAAGAAA
1150 ACATAGGAATTTATAAAGAACTAGAAGATTTGATAGAGAAAAATGAAAATT
1201 TAGATGATTTAGATGAAGGAATAGAAAAATCATCAGAAGAATTATCTGAAG
1252 AAAAAATAAAAAAAGGAAAGAAATATGAAAAAACAAAGGATAATAATTTTA
1303 AACCAAATGATAAAAGTTTGTATGATGAGCATATTAAAAAAATATAAAAATG
1354 ATAAGCAGGTTAATAAGGAAAAGGAAAAATTCATAAAATCATTGTTTCATA
1405 TATTTGACGGAGACAATGAAATTTTACAGATCGTGGATGAGTTATCTGAAG
1456 ATATAACTAAATATTTTATGAAACTA<u>TAA</u> (stop)  (SEQ ID NO: 39)
```

(NON-CODING 3' END)

```
1485 AAGGTTATATATTT 1498
```

FIGURE 8

LSA.3'.ALL -> 1-phase Translation

DNA sequence 1496 b.p.   (SEQ ID NO: 40) CAAGAACAACAA ... (SEQ ID NO: 41) GGTTATATATTT   linear

(SEQ ID NO: 42)
(SEQ ID NO: 43)

```
  1 /   1                                                      31 /  11
CAA GAA CAA CAA AGC GAT CTA GAA CAA AGA CGT GCT AAA GAA AAG TTG CAA GAA CAA
gln glu gln gln ser asp leu glu gln arg arg ala lys glu lys leu gln glu gln
 61 /  21                                                      91 /  31
CAA AGC GAT TTA GAA CAA CAA GAT AGA CTT GCT AAA GAA AAG TTA CAA GAG CAG CAA AGC GAT
gln ser asp leu glu gln gln asp arg leu ala lys glu lys leu gln glu gln ser asp
121 /  41                                                     151 /  51
TTA GAA CAA GAG AGA CTT GCT AAA GAA AAG TTG CAA GAG CAG CAA AGC GAT CTA GAA CAA
leu glu gln glu arg leu ala lys glu lys leu gln glu gln ser asp leu glu gln
181 /  61                                                     211 /  71
GAG AGA CGT GCT AAA GAA AAG TTG CAA GAG CAG CAA CAA AGC GAT TTA GAA CAA CGT
glu arg arg ala lys glu lys leu gln glu gln gln ser asp leu glu gln arg
241 /  81                                                     271 /  91
GCT AAA GAA AAG TTG CAA GAG CAG CAA CAA AGC GAT TTA GAA CAA GAT CTA GAA CAA
ala lys glu lys leu gln glu gln gln ser asp leu glu gln asp leu glu gln
301 / 101                                                     331 / 111
AAG TTA CAA GAG CAG CAA AGC GAT TTA GAA CAA GAG CAA AGA CGT GCT AAA GAA TTG CAA
lys leu gln glu gln ser asp leu glu gln glu gln arg arg ala lys glu leu gln
```

FIGURE 9A

```
1261  /  421                 1291  /  431
AAA GGA AAG AAA TAT GAA      AAA ACA AAG GAT AAT AAT TTT AAA CCA AAT GAT AAA AGT TTG
lys gly lys lys tyr glu      lys thr lys asp asn asn phe lys pro asn asp lys ser leu
1321  /  441                 1351  /  451
TAT GAT GAG CAT ATT AAA      AAA TAT AAA AAT GAT AAG CAG GTT AAT AAG GAA GAA AAA
tyr asp glu his ile lys      lys tyr lys asn asp lys gln val asn lys glu lys lys
1381  /  461                 1411  /  471
TTC ATA AAA TCA TTG TTT      CAT ATA TTT GAC GGA GAC AAT GAA ATT TTA CAG ATC GTG GAT
phe ile lys ser leu phe      his ile phe asp gly asp asn glu ile leu gln ile val asp
1441  /  481                 1471  /  491
GAG TTA TCT GAA GAT ATA      ACT AAA TAT TTT ATG AAA CTA TAA AAG GTT ATA TAT
glu leu ser glu asp ile      thr lys tyr phe met lys leu [OCH lys val phe tyr]
```

✗ Strike from Fig

FIGURE 9D

LSN.3'STOP -> 1-phase Translation

(SEQ ID NO: 45)

(SEQ ID NO: 44)

DNA sequence   1482 b.p.   CAAGAACAACAA ... ATGAAACTATAA   linear

```
1 / 1
(SEQ ID NO: 46) CAA GAA CAA AGC GAT CTA GAA CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA GAA
                31 / 11
(SEQ ID NO: 47) gln glu gln ser asp leu glu gln glu arg arg ala lys glu lys leu gln glu gln
                61 / 21
CAA AGC GAT TTA GAA CAA GAT AGA CTT GCT AAA GAA AAG TTA CAA GAG CAG CAA AGC GAT
gln ser asp leu glu gln asp arg leu ala lys glu lys leu gln glu gln gln ser asp
121 / 41
TTA GAA CAA GAG AGA CTT GCT AAA GAA AAG TTG CAA GAA CAA CAA GAT CTA GAA CAA
leu glu gln glu arg leu ala lys glu lys leu gln glu gln gln ser asp leu glu gln
181 / 61
GAG AGA CGT GCT AAA GAA AAG TTG CAA GAA CAA CAA AGC GAT TTA GAA CAA GAG AGA CGT
glu arg arg ala lys glu lys leu gln glu gln gln ser asp leu glu gln glu arg arg
241 / 81
GCT AAA GAA AAG TTG CAA GAA CAA CAA GAT AGA CTT GCT AAA GAA
ala lys glu lys leu gln glu gln gln ser asp arg leu ala lys glu
301 / 101
AAG TTA GAA CAA GAG CAG CAA GAG AGA CGT GCT AAA GAA AAG TTG CAA
lys leu glu gln glu gln gln glu arg arg ala lys glu lys leu gln
```

FIGURE 10A

```
1321  /   441              1351  /  451
TAT GAT GAG CAT ATT AAA AAA TAT AAA AAT GAT AAG CAG GTT AAT AAG GAA AAG GAA AAA
tyr asp glu his ile lys lys tyr lys asn asp lys gln val asn lys glu lys glu lys
1381  /   461              1411  /  471
TTC ATA AAA TCA TTG TTT CAT ATA TTT GAC GGA GAC AAT GAA ATT TTA CAG ATC GTG GAT
phe ile lys ser leu phe his ile phe asp gly asp asn glu ile leu gln ile val asp
1441  /   481              1471  /  491
GAG TTA TCT GAA GAT ATA ACT AAA TAT TTT ATG AAA CTA TAA AAG GTT ATA TAT
glu leu ser glu asp ile thr lys tyr phe met lys leu [och lys val ile tyr]

                                              *DELETE FROM FIGURE
```

FIGURE 10D